

# COMPUTER SYSTEM CONTAINING OPTICAL FIBER-CHANNEL CONNECTION TYPE STORAGE MECHANISM

Publication number: JP10293633

Publication date: 1998-11-04

Inventor: SKIDMORE ALAN E; SCHULTZ STEPHEN M;  
SCHOLHAMER GEORGE J; THOMPSON MARK J;  
RUSHTON TOD D; FREEMAN EUGENE E;  
ALEXANDER DENNIS J; CALLISON RYAN A; GRIEFF  
THOMAS W; SABOTTA MICHAEL L; GRANT DAVID L;  
MCCARTY JAMES F; SCHNEIDER RANDY D;  
GALLOWAY WILLIAM C

Applicant: COMPAQ COMPUTER CORP

Classification:

- International: G06F3/00; G06F3/00; (IPC1-7): G06F3/00

- European:

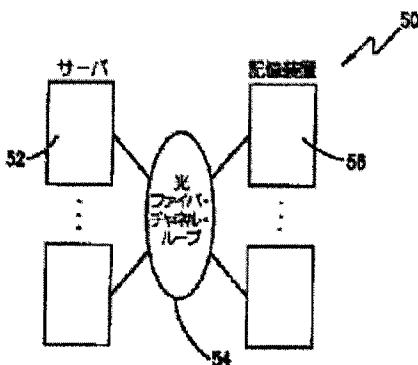
Application number: JP19980043374 19980225

Priority number(s): US19970805281 19970225

[Report a data error here](#)

## Abstract of JP10293633

**PROBLEM TO BE SOLVED:** To provide a storage mechanism architecture which has the excellent performance. **SOLUTION:** In this system or a network, at least a server 52 is connected to at least a storage 56 via an optical fiber/channel loop 54. In such a constitution, the redundancy is secured for a connection mechanism and a high data rate is attained. Then plural operating systems are available and a system containing a hot plugging function is attained. Furthermore, many devices such as the servers 52 and storages 56 or the devices related to other systems can be connected to the loop 54. The distances among these connectable devices can be set at 10 kilometers or more. Even in such a case, the communication is possible among the devices at a high data rate with use of the loop 54.



(19)日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-293633

(43)公開日 平成10年(1998)11月4日

(51)Int.Cl.<sup>6</sup>  
G 0 6 F 3/00

識別記号

F I  
G 0 6 F 3/00

E

審査請求 未請求 請求項の数19 O.L (全 10 頁)

(21)出願番号 特願平10-43374

(22)出願日 平成10年(1998)2月25日

(31)優先権主張番号 8 0 5 2 8 1

(32)優先日 1997年2月25日

(33)優先権主張国 米国(US)

(71)出願人 591030868  
コンパック・コンピューター・コーポレーション  
COMPAQ COMPUTER CORPORATION  
アメリカ合衆国テキサス州77070, ヒューストン, ステイト・ハイウェイ 249,  
20555(72)発明者 アラン・イー・スキッドモア  
アメリカ合衆国テキサス州77070, ヒューストン, オーク・パーク・ドライブ  
12315

(74)代理人 弁理士 松本 一夫 (外5名)

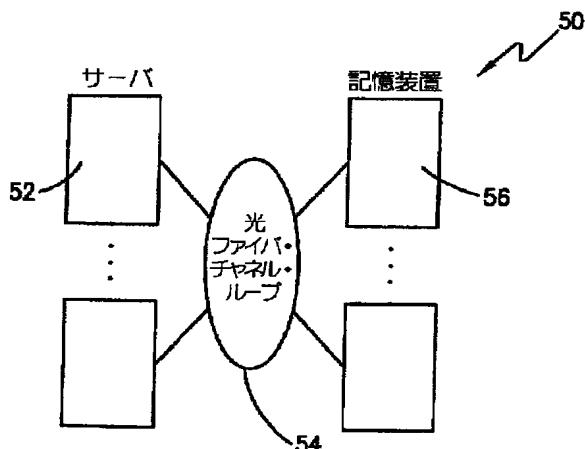
最終頁に続く

(54)【発明の名称】光ファイバ・チャネル接続式記憶機構を備えたコンピュータ・システム

## (57)【要約】

【課題】 優れた性能を有する記憶機構アーキテクチャを提供する。

【解決手段】 光ファイバ・チャネル54を介して少なくとも1台のサーバ52を少なくとも1台の記憶装置56に接続するようにしたシステムないしネットワークである。このシステムによれば、接続機構に冗長性を持たせることができ、データ・レートを高速化することができ、複数のオペレーティング・システムを使用することができ、また、ホット・プラギング機能を備えたシステムとすることができます。更に、このシステムによれば、多数のデバイスを光ファイバ・チャネルに接続することができる。接続するデバイスは、サーバ、記憶装置、或いはその他のシステムに関連した機器であり、それらデバイス間の離隔距離を10キロメートル以上にすることができる、それだけ離隔させた場合でも、ファイバ・チャネルを使用しているため高速のデータ・レートで通信を行うことができる。



## 【特許請求の範囲】

【請求項1】 コンピュータ・システムにおいて、光ファイバ・チャネルから成る相互接続機構と、前記光ファイバ・チャネルに接続されたサーバと、前記光ファイバ・チャネルに接続された記憶装置とを備えたことを特徴とするコンピュータ・システム。

【請求項2】 請求項1記載のコンピュータ・システムにおいて、前記サーバと前記光ファイバ・チャネルとがホスト・バス・アダプタを介して接続されていることを特徴とするコンピュータ・システム。

【請求項3】 請求項1記載のコンピュータ・システムにおいて、前記サーバと前記光ファイバ・チャネルとがメディア・モジュールを介して接続されていることを特徴とするコンピュータ・システム。

【請求項4】 請求項1記載のコンピュータ・システムにおいて、前記光ファイバ・チャネルと前記記憶装置とがアレイ・コントローラを介して接続されていることを特徴とするコンピュータ・システム。

【請求項5】 請求項1記載のコンピュータ・システムにおいて、該システムは、前記光ファイバ・チャネルに複数台の前記サーバが接続されていることを特徴とするコンピュータ・システム。

【請求項6】 請求項1記載のコンピュータ・システムにおいて、前記光ファイバ・チャネルに複数台の前記記憶装置が接続されていることを特徴とするコンピュータ・システム。

【請求項7】 請求項1記載のコンピュータ・システムにおいて、前記光ファイバ・チャネルに複数台の前記サーバと複数台の前記記憶装置とが接続されていることを特徴とするコンピュータ・システム。

【請求項8】 請求項1記載のコンピュータ・システムにおいて、前記光ファイバ・チャネルに光ファイバ・チャネル・ループがカスケード接続されていることを特徴とするコンピュータ・システム。

【請求項9】 請求項1記載のコンピュータ・システムにおいて、該システムは更に冗長光ファイバ・チャネルを備えており、該冗長光ファイバ・チャネルが前記サーバと前記記憶装置とに接続されていることを特徴とするコンピュータ・システム。

【請求項10】 請求項1記載のコンピュータ・システムにおいて、該システムは更に第2のサーバを備えており、該第2のサーバは前記光ファイバ・チャネルに接続されていて前記サーバとは異なったオペレーティング・システムで動作するものであることを特徴とするコンピュータ・システム。

【請求項11】 請求項1記載のコンピュータ・システムにおいて、該システムは更に第2の記憶装置を備えており、該第2の記憶装置は第2の記憶手段を構成しており、前記記憶装置は第1の記憶手段を構成していることを特徴とするコンピュータ・システム。

【請求項12】 請求項1記載のコンピュータ・システムにおいて、前記コンピュータ・システムに電源が入った状態で前記光ファイバ・チャネルにホット・プラギング方式でデバイスを接続できるようにしてあることを特徴とするコンピュータ・システム。

【請求項13】 光ファイバ・チャネル接続式記憶機構アーキテクチャにおいて、光ファイバ・チャネルから成る相互接続機構と、前記光ファイバ・チャネルに接続された少なくとも1台のサーバと、前記光ファイバ・チャネルに接続された少なくとも1台のデータ記憶装置とを備えたことを特徴とする光ファイバ・チャネル接続式記憶機構アーキテクチャ。

【請求項14】 請求項13記載の光ファイバ・チャネル接続式記憶機構アーキテクチャにおいて、前記光ファイバ・チャネルにホット・プラギング方式でデバイスを追加して接続できるようにしてあることを特徴とする光ファイバ・チャネル接続式記憶機構アーキテクチャ。

【請求項15】 請求項13記載の光ファイバ・チャネル接続式記憶機構アーキテクチャにおいて、前記少なくとも1台のサーバと前記少なくとも1台の記憶装置との間で100メガバイト/秒以上のデータ・レートでデータを転送できるようにしてあることを特徴とする光ファイバ・チャネル接続式記憶機構アーキテクチャ。

【請求項16】 請求項13記載の光ファイバ・チャネル接続式記憶機構アーキテクチャにおいて、前記少なくとも1台のサーバと前記少なくとも1台の記憶装置との間の離隔距離を6メートル以上にしてあることを特徴とする光ファイバ・チャネル接続式記憶機構アーキテクチャ。

【請求項17】 請求項13記載の光ファイバ・チャネル接続式記憶機構アーキテクチャにおいて、前記少なくとも1台のサーバが複数台のサーバであり、それら複数台のサーバのうちの少なくとも2台のサーバが互いに異なったオペレーティング・システムを使用していることを特徴とする光ファイバ・チャネル接続式記憶機構アーキテクチャ。

【請求項18】 請求項13記載の光ファイバ・チャネル接続式記憶機構アーキテクチャにおいて、該アーキテクチャは更に第2の光ファイバ・チャネルを備えており、該第2の光ファイバ・チャネルが前記少なくとも1台のサーバと前記少なくとも1台の記憶装置とに接続されていることを特徴とする光ファイバ・チャネル接続式記憶機構アーキテクチャ。

【請求項19】 請求項13記載の光ファイバ・チャネル接続式記憶機構アーキテクチャにおいて、該アーキテクチャは更に第2の光ファイバ・チャネルを備えており、該第2の光ファイバ・チャネルを前記光ファイバ・チャネルに接続することで該第2の光ファイバ・チャネルに更にサーバ及び記憶装置を追加して接続できるよう

にしてあることを特徴とする光ファイバ・チャネル接続式記憶機構アーキテクチャ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、光ファイバ・チャネルから成る接続機構を介してコンピュータ・システムをデータ記憶装置に接続するようにしたシステムないしネットワークのアーキテクチャに関する。より詳細に言えば、本発明は、光ファイバ・チャネル仲裁ループから成るネットワークを介して1台ないし複数台のネットワーク・サーバを1台ないし複数台の記憶装置に接続するようにした接続機構の構成に関する。

【0002】

【従来の技術】近年、コンピュータ・ネットワークにおける大容量記憶装置の重要性がますます増大している。実際に、ある種のシステムに装備されている大容量記憶装置では、ギガバイト単位のデータがより多く格納されるために、その記憶空間がとどまることなく増大している。今日のコンピュータ・ネットワークでは、大容量記憶装置が十分な記憶容量を有することに加えて、その大容量記憶装置に格納されているデータへのアクセスが容易であることや、その大容量記憶装置に対するデータの読み書きを行う際のデータ・レートが高速であることも要求されている。実際に、大容量記憶システムは、そのデータ・レートが高速であるほど、また、エラー・フリ一性能が高いほど、優れたシステムであるといえる。

【0003】現在、コンピュータの業界において、コンピュータ・システムと大容量記憶装置とを接続するための方式として広く採用されているのはSCSI規格である。SCSIバスは、約20メガバイト／秒の通信速度で通信を行うことができる。図1は従来の構成を示したものであり、この構成においては、ディスク記憶装置10とネットワーク・サーバ12とがSCSIバス14を介して互いに接続されている。

【0004】

【発明が解決しようとする課題】SCSIバスには、コンピュータの業界が現在熱心に取り組んでいる様々な問題や短所が付随している。その一例を挙げるならば、SCSIバスから成る接続機構のデータ・レートは、サーバにとっては不十分なものである。即ち、SCSIバスのデータ・レートは、現在のネットワーク・サーバの通信速度及び機能を十分に活用するためには遅すぎ、また、SCSIバスは、現在のネットワーク・サーバに対応した十分に大きな記憶容量を有するディスク記憶装置にアドレスするために必要な性能を備えているとはいえない。SCSIバスで構成した接続機構を用いた場合には、大容量記憶機構のアーキテクチャにおける相互接続機構に冗長性を持たせることができないため、サーバまたは記憶装置の一部に故障が発生したときに、ユーザが記憶装置のデータにアクセスできなくなる恐れがある。

【0005】また、SCSIバスで構成した接続機構を用いた場合には、1台の記憶装置に格納されているデータを、種類の異なるサーバが共用できるようになることが容易でない。更に、SCSIバスを用いたシステムは、ホット・プラギング可能なシステムとすることが困難である。即ち、システムが動作中であって、データがネットワーク・バス(SCSIバス)上を転送されているときに、サーバや記憶ユニット等のデバイスをそのSCSIバスに「プラグ接続」によって追加して接続できるようになることが容易でない。更にSCSIバスの配線用ケーブルは、重く、かさばり、太く、屈曲させにくい。

【0006】従って、コンピュータ・サーバ及びコンピュータ・ネットワークに用いる大容量記憶装置の分野においては、上述したSCSIバスの様々な短所を改善することが切望されている。本発明に係る光ファイバ・チャネル接続式記憶機構アーキテクチャは、従来のSCSIバスを用いた記憶機構アーキテクチャに付随していた上記した問題及び短所を克服することを目的としたものである。

【0007】

【課題を解決するための手段】上記した目的を達成するために、本発明においては、SCSIバスを使用することをやめて、その代わりに光ファイバ・チャネル接続式記憶機構アーキテクチャを採用することを特徴としている。本発明にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャは、サーバと記憶装置との間で行われるデータ読み書きの速度を、100メガバイト／秒以上(SCSIバスを用いた場合のデータ読み書き速度の5倍以上)に向上させることができる。また、1台のサーバを複数台の記憶装置に接続することができるため、サーバがアクセスすることのできる記憶容量が増大する。また、複数台の記憶装置を複数台のサーバに接続することができるため、サーバ・クラスタないしサーバ・ネットワークを形成している複数台のサーバの間でのデータの共用が可能となる。本発明にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャは更に、「ホット・プラギング」機能を備えたものとすることができ、それによって、光ファイバ・チャネルの動作中に、サーバや、記憶装置や、その他のデバイスを、その光ファイバ・チャネルに着脱することが可能になる。

【0008】

【発明の実施の形態】これより本発明の具体的な実施の形態について、図面を参照して詳細に説明する。本発明の好適な実施例の光ファイバ・チャネル接続式記憶機構アーキテクチャは、将来開発されるコンピュータ・システムの一部を構成する要素として利用し得るものである。また、この光ファイバ・チャネル接続式記憶機構アーキテクチャを用いれば、冗長性を有することによって障害及び転送時の誤りを回復することのできる機能を備

えた大記憶容量のデータ記憶及び取出機構を構成することができる。更に、この光ファイバ・チャネル接続式記憶機構アーキテクチャは、データ・バスを介して送信するデータの送信距離を大きく取れるという点において、それに、現在の外付式の記憶機構サブシステムを容易に使用できるという点において、大きな融通性を提供するものである。

【0009】本発明は、光ファイバ・チャネルを使用するようにしたものである。光ファイバ・チャネルは、様々なトポロジーで構成することのできるマルチ・トポロジーの物理転送チャネルであり、SCSI、IP、それにIPX等の、複数のアッパー・レイヤ・プロトコルに従って転送を行う共通物理レイヤとして機能するものである。また、光ファイバ・チャネルを使用することで、仲裁ループを低成本で構築することができる。光ファイバ・チャネルは、ディジタル・ビデオ・システムやディジタル撮像システム等において要求される、大きなブロック・サイズのデータ・ブロックによるデータ転送を取り扱う場合に理想的なものである。

【0010】光ファイバ・チャネルは、バス幅の広いSCSIバスと比べて、データ転送速度に関する大きな利点を提供するものである（データ・レートをSCSIバスの少なくとも5倍にすることができる）。光ファイバ・チャネルを用いることで、複数の種類のプロトコルに対応することのできる非常に高性能の相互接続機構を構築することができる。これらの利点は、ディスク・ドライブ、サーバ・ネットワーク、ワークステーション、及びパーソナル・コンピュータのコストを引き下げるのに役立ち、また、それらの装置においてスーパー・コンピューティングを行うことを可能にするものである。

【0011】光ファイバ・チャネルは構成が比較的簡明である。光ファイバ・チャネルは、基本的に、ポイントツーポイント（ポイント対ポイント）接続を提供する接続機構であるが、ただしそれを、複数のポイントツーポイント接続を提供する仲裁ループ（論理ループ）とすることで、複数個のノードにも対応できるようになる。図示の例では、仲裁ループのプロトコルによって127個の物理ノードへのアドレスを可能にしている。また、この仲裁ループにおいては、通信経路を確立するための中央交換装置を使用せずに、各ノードが仲裁プロトコルに従って動作することで、チャネル上で通信が行われるようになっている。即ち、あるノードが通信を行う権利を獲得したならば、そのノードは通信先ノードとの間の接続経路を確立し、その接続経路が確立したならば、それら2つの当事者ノードの間で通信が開始される。この通信は、全二重方式で行われ、また、1.0625ギガビット/秒のデータ転送速度で行われる。それら2つの当事者ノード以外のその他のノードはいずれもリピータ即ち中継器として動作する。光ファイバ・チャネル・ループとして構成した接続機構を用いることでサーバと記憶装

置との間の関係が融通性に富んだものとなり、それは、（SCSIバスを用いた場合のケーブルの配線に必要とされる）厳密な1対1の直接的なケーブルの関係が、最早必要ではなくなるからである。

【0012】図2には、本発明の好適な実施の形態にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャ50の基本的なブロック図が示されている。光ファイバ・チャネル・ループ54には、1台または複数台のサーバ52が接続される。更にこの光ファイバ・チャネル・ループ54には、1台または複数台の記憶装置56が接続される。そして、この光ファイバ・チャネル・ループ54によってマッピングされる関係をもって、その1台または複数台のサーバ52が、その1台または複数台の記憶装置56に接続されることになる。

【0013】図2に示した光ファイバ・チャネル接続記憶機構アーキテクチャの構成は、幾つもの利点を有する。第1に、各サーバ52からアクセス可能な記憶装置56の台数が（即ち、サーバ52の1つのスロットに接続できる記憶装置の台数が）増大するため、光ファイバ・チャネル・ループ54に接続されている各サーバ52がアクセスすることのできる記憶容量が増大する。第2に、光ファイバ・チャネル・ループ54を使用することで、光ファイバ・チャネル・ループ以前の世代の方式に付随していた制約を受けることなく、記憶装置56を共用することが可能になるため、データの可用性が向上する。第3に、この実施の形態によれば「ホット・グロース（hot growth）」が可能となる。これは、最初は僅かな台数のサーバと記憶装置とで光ファイバ・チャネル接続記憶機構アーキテクチャを構成しておき、

後に、システムに電源が入っていて光ファイバ・チャネル・ループ54上をデータが転送されている状態で、記憶装置56を光ファイバ・チャネル・ループ54に「ホット・プラギング方式」または「ホット・プラグ方式」と称される方式で接続することによって、そのシステムを拡張できるということである。ホット・プラギング方式で接続がなされたときに光ファイバ・チャネル・ループ54上をデータが転送されていた場合には、システムは誤り回復サイクルを実行して、「壊れた」データ・パケットの一部である可能性のあるデータに対して訂正を施す。従って、ここでいうホット・プラギングには、単にシステムに電源が入っている状態で接続を行うことばかりでなく、システムがデータ・フレームを転送している状態で接続を行うことも含まれている。

【0014】この実施例においてホット・プラギングが可能であることによるもう1つの利点として、媒体の交換が可能なシステムを構築できることがある。これは例えば、最初は、長さが約300～500メートルの光ファイバ・チャネル・ループとして構成したマルチ・モードの光ファイバから成る接続機構を備えたアーキテクチャを構築する。そして、ユーザがそのネットワークをア

アップグレードしたいと考えたときに、その接続機構を、データを10キロメートルもの遠くまで転送することのできるシングル・モードの光ファイバから成る接続機構に交換すればよい。また、この実施の形態にかかるアーキテクチャによれば、当業者は、アップグレードしたハードウェアを容易に付加する（例えばプラグを差し込むことで接続する）ことができ、その際に記憶装置やサーバのソフトウェアに変更を加える必要がない。従って、ホット・プラギング機能を備えたことによって、この実施の形態は非常に融通性に富み、またアップグレードの容易なものとなっている。この実施の形態にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの第4の利点は、光ファイバ・チャネルのデータ転送性能が優れているため、記憶装置に対するデータの読み書き速度が向上することである。例えば、SCSIバス・システム（図1）のデータ・レートは、約20メガバイト／秒である。これまでに、SCSIバスのデータ転送速度を向上させるための方法が幾つも提案されているが、それら方法によって達成される転送速度の向上は僅かなものでしかなかった。

【0015】これに対して、この実施の形態において使用している光ファイバ・チャネル・ループは、100メガバイト／秒以上のデータ・レートを提供する能力を備えている。この実施の形態にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャによって得られるデータ・レートの向上は、SCSI方式と比べれば、まさに大きな進歩というべきものである。第5の利点は、本発明を採用することによって記憶装置の増設が容易になるということである。即ち、当業者であれば、本発明を採用することで、同一の相互接続機構に接続した一次記憶装置と二次記憶装置とを統合して、それら一次記憶装置と二次記憶装置とが、1台のサーバのデータベースから、或いは1つのサーバ・クラスタのデータ・ベースからは、单一の大きな記憶装置であるかのように見えるようになることが比較的容易になる。

【0016】この実施例にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの第6の重要な局面は、光ファイバ・チャネル・ループに接続した複数のデバイスを互いにどれだけ離すことができるかという、デバイス間の離隔距離に関するものである。即ち、従来のSCSIバスを用いたシステムでは、SCSIバスに課されている制約のために、どのサーバないし記憶装置の間の離隔距離も数メートル以内にとどめなければならなかつた。これに対して、光ファイバ・チャネル・ループに、この実施の形態に使用しているサーバ及び記憶装置というデバイスを接続した場合には、2台のデバイス間のポイントツーポイントの離隔距離を、シングル・モードの光ファイバを使用したときに、約10キロメートルもの長い距離にすることができる。この実施の形態のアーキテクチャによれば、ユーザはデータないしサーバの災害

復旧や、ミラー・データないしミラー・サーバの遠隔作成、このような大きな離隔距離でもって行うことができる。更に加えて、例えば、1つの大学敷地内ないしは企業敷地内に離ればなれに建てられている複数の建物に分散して設置されているサーバや記憶装置の全てを容易に接続することができる。

【0017】この実施例にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの第7の局面は、そのケーブルが光ファイバであり、その光ファイバ・チャネル・ループのプロトコルとして、パラレル・バス・プロトコルではなくシリアル・プロトコルが使用されていることである。この構成によれば、データの論理転送のために必要なデータ信号線は基本的に1本でよいため、従来のSCSIバスのパラレル・ケーブルと比べて接続ケーブルを非常に細くすることができる。このようにケーブルが細いということも、この実施の形態における重要な特質のうちの1つである。即ち、ユーザが小規模な企業であって、机上に設置するコンピュータを使用している場合には、ケーブルの太さが問題となることは殆どないと思われるが、一方、ユーザが例えば大規模な企業であって、ラックに多数のサーバを収容して使用している場合には、多数のケーブルが混み合ってラック内を縦横に走ることになり、それによって問題が生じるおそれがある。従って、そのような場合には、ケーブル自体が細いこと、ケーブルの接続部を小型化できること、それにシステムを（空間的に）分散させることができるとの能力を備えていることが大きな利点となる。

【0018】この実施例にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの第8の好適な局面は、種類の異なるプラットフォームの間で記憶装置を1つにまとめることができる。即ち、製造メーカーが異なる複数台のサーバを使用する場合でも、本発明にかかるアーキテクチャによれば、それらサーバや、記憶装置や、夫々に異なったプロトコルで動作しているサブシステムを、单一の光ファイバ・チャネル接続式記憶アーキテクチャに接続することで、通信及び動作を効率的に行う能力をそのまま維持することができる。

【0019】本発明は、更にその他の様々な実施の形態としても構成し得るものであり、それら実施の形態について以下に説明する。それら実施の形態には、光ファイバ・チャネル接続式記憶機構の基本構成、共用冗長経路記憶機構の構成、統合二次記憶機構の構成、マルチブル・ネットワーク・オペレーティング・システム共用記憶機構の構成、それに異種記憶システム共用機構の構成が含まれる。尚、1つの構成の上に更に別の構成を重ねて構築したり、異なる構成を組み合わせて用いることも可能である。

【0020】図3は、本発明にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの第2の実施例を示した図である。この第2の実施例は、光ファイバ・チャ

ネル接続式記憶機構の基本構成60を示したものである。1台だけ装備されているサーバ62は、ホスト・バス・アダプタ回路64を備えており、メディア・モジュール66は、サーバ62を光ファイバ・チャネル・ループ68に接続されている。メディア・モジュール66は、サーバ62を複数台の記憶ユニット(記憶装置)70が接続されており、それら記憶ユニット(記憶装置)70の各々は、メディア・モジュール66を介して光ファイバ・チャネル・ループ68に接続されている。メディア・モジュール66はアレイ・コントローラ72に接続されている。アレイ・コントローラ72は、この記憶ユニット70に収容される様々なディスク・ドライブ74を制御することができるよう設計されている。この実施の形態によれば、従来のSCSI方式を採用したシステムや、その他の従来のシステムと比べて、サーバ62がアクセスすることのできる記憶容量が増大する。

【0021】記憶ユニット70に収容するディスク・ドライブ74の台数の好適例は、例えば8台である。各ディスク・ドライブ74の記憶容量は例えば9ギガバイトである。この場合、1台の記憶ユニット70の記憶容量は、ディスク・ドライブ1台あたり9ギガバイトの8倍で72ギガバイトになる。ここで、光ファイバ・チャネル・ループ68に接続した記憶ユニット(記憶装置)70の台数が6台であるとするならば、この光ファイバ・チャネル接続式記憶機構アーキテクチャに接続されている各サーバ62が利用できる記憶容量は、記憶ユニット1台あたり72ギガバイトの6倍で432ギガバイトになる。この実施の形態では、記憶ユニット70に収容されているディスク・ドライブ74は、SCSI規格のハード・ディスク・ドライブである。尚、記憶ユニット70は様々な構成のものとすることができる、また記憶ユニット70の記憶容量も、その中に収容した記憶装置(ドライブ)に応じて、以上に例示したものより大容量にもなることもあれば小容量になることもある。

【0022】図3に示した第2の実施例にかかる光ファイバ・チャネル・ループの構成は、バイパス・ファイバ・ループ・ハブ68を1つだけ備えており、そこにサーバ62が1台だけ接続されている。ホスト・バス・アダプタ64の好適な具体例は、例えば、PCIと光ファイバ・チャネルとを接続するホスト・バス・アダプタである。ホスト・バス・アダプタ64は、ホスト・サーバのPCIスロットまたはEISAスロットに挿入するプリント回路板として構成することができる。ホスト・バス・アダプタ64は、ホスト・サーバのPCIバスまたはEISAバスから光ファイバ・チャネル・ループ68へのマッピングを行う機能を果たすものである。また基本的に、このホスト・バス・アダプタ64は、ホスト・サーバのPCIバスまたはEISAバスから受け取ったデ

ータを光ファイバ・チャネル・ループ68のプロトコルに適合させるための、プロトコル翻訳を行うものである。

【0023】記憶ユニット70の中に装備した図示例のアレイ・コントローラ72は、光ファイバ・チャネルとSCSI規格のデバイスとを接続するためのアレイ・コントローラである。このアレイ・コントローラ72の好適な具体例は、マイクロプロセッサを搭載したプリント回路板として構成したものである。アレイ・コントローラ72は、複数台の記憶装置から成る記憶装置アレイを制御するものである。なお、ここでいう記憶装置とは、例えば、ディスク・ドライブ、テープ・ドライブ、それに光読み書きドライブ等である。アレイ・コントローラ72は、フォールト・トレランス機能を提供するようにプログラムされている。メディア・モジュール66は、短波長光ファイバのためのメディア変換デバイスであり、光ファイバ信号を電気信号に変換する機能と、逆に電気信号を光ファイバ信号に変換する機能とを果たすものである。

10 【0024】図4は、本発明にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの更に別の実施例を示した図である。この実施の形態は、図3に示した実施例を拡張したものであり、その拡張の内容は、記憶機構の1つの光ファイバ・チャネル・ループに複数台のサーバを接続したことである。この図4の実施例では、2台以上のサーバが同一のデータにアクセスすることができる。それゆえこの実施例における記憶機構は、共用記憶機構である。従ってこの実施例は、共用記憶機構の構成を採用したものである。2台以上のサーバが同一のデータにアクセスできるようにしたこの利点は、それによってフェイル・オーバ・プロテクション機能が得られるということであり、このフェイル・オーバ・プロテクション機能とは、サーバのうちの1台が故障した場合でもデータにアクセスできるようにする機能のことである。

20 【0025】図4に示した形態では、複数台のサーバ62が、夫々にホスト・バス・アダプタ64とメディア・モジュール66とを介して光ファイバ・チャネル・ループ68に接続されている。更に、光ファイバ・チャネル・ループ68は、夫々にメディア・モジュール66を介して複数台の記憶ユニット70に接続されている。記憶ユニット70はアレイ・コントローラ72を備えており、このアレイ・コントローラ72は、複数台のディスク・ドライブ74に対するデータの流れを制御している。

30 【0026】図4を参照すると、光ファイバ・チャネル・ループ68をカスケード接続できることが示されている。カスケード接続する光ファイバ・チャネル・ループは、独立した光ファイバ・チャネル仲裁ループであることを必ずしも必要としない。カスケード接続して付加した光ファイバ・チャネル・ループは、元からある光ファ

40 40 【0026】図4を参照すると、光ファイバ・チャネル・ループ68をカスケード接続できることが示されている。カスケード接続する光ファイバ・チャネル・ループは、独立した光ファイバ・チャネル仲裁ループであることを必ずしも必要としない。カスケード接続して付加した光ファイバ・チャネル・ループは、元からある光ファ

11  
イバ・チャネル・ループを拡張することになり、ここでいう拡張とは、より多くのポートが使用可能になるという意味である。幾つもの光ファイバ・チャネル・ループをカスケード接続する場合には、交換装置等のデバイスを装備するようにしてよく、そうすることによって、それら光ファイバ・チャネル・ループの性能及び管理容易性を向上させることができる。更に具体的に説明すると、光ファイバ・チャネル・ループ68は複数個のノードを備えている。ここで、光ファイバ・チャネル・ループ68が備えているノードの個数が7個であるものとすれば、それらのうちの6個のノードを、サーバ、記憶装置、ないしはその他の周辺機器に接続し、残った7番目のノードを第2の光ファイバ・チャネル・ループ68Aにカスケード接続するために使用すればよい。更に、第2の光ファイバ・チャネル・ループ68Aも同じく7個のノードを備えているものとすれば、この第2の光ファイバ・チャネル・ループ68Aの残りの6個のノードを、更にその他の光ファイバ・チャネル・ループにカスケード接続したり、或いは、サーバ、記憶装置、ないしはその他の周辺機器に接続するために使用すればよい。

【0027】図5は、二重光ファイバ・チャネル・ループを装備することで、共用冗長経路記憶機構を構成した実施例を示した図である。この構成とする場合には、例えば、第1の光ファイバ・チャネル仲裁ループから（信号的にも物理的にも）独立した第2の光ファイバ・チャネル仲裁ループで、冗長ループを構成するようにすればよい。各記憶ユニット70は、第1のアレイ・コントローラ72に加えて更に第2のアレイ・コントローラ72Aを装備しており、また、各サーバ・ホスト62は、第1のホスト・バス・アダプタ64に加えて更に第2のホスト・バス・アダプタ64Aを装備している。この実施の形態では、二重光ファイバ・チャネル・ループを備えたことによって実用性が大いに高められている。ここで可用性が高められているというのは、あるサーバ62にログインしたユーザが、そのサーバ62や、アダプタ、I/O経路、記憶ユニット、電源、ファン、コントローラ等に故障が発生した場合でも、確実にデータにアクセスできるという意味である。また、この構成においては、単に複数の記憶ユニットのデータを複数のサーバの間で共用できるばかりでなく、その共用を冗長性をもって行えるようになっている。この冗長性ということも、図5の実施の形態にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャによって得られる特徴的構成の1つであり利点の1つである。

【0028】I/O経路が完全冗長性を備えている（即ち、2つの光ファイバ・チャネル・ループを備えている）ため、一方のI/O経路に障害が発生しても、他方のI/O経路が提供する冗長性によってその障害の影響を被らずに済む。この実施例は、先に説明した幾つかの実施例を内包する上位集合に相当する実施例であるとい

12  
える。図5の実施の形態のアーキテクチャは、共用冗長経路記憶機構を構成するものであって、フェイル・オーバ機能を備えた完全冗長性を有する2つのI/O経路を提供するものである。また、この実施の形態においても、光ファイバ・チャネル・ループはホット・ブリッジングが可能なように構成されており、この実施の形態は、先に説明した幾つかの実施の形態にかかるアーキテクチャの特徴を全て備えている。

【0029】図6に示したのは、本発明にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの更に別の実施例である。この実施例は、統合二次記憶機構の構成をアーキテクチャに組み込んだものである。この図6のアーキテクチャは、1つの相違点を除いて図5のアーキテクチャと基本的に同じものであり、その相違点とは、光ファイバ・チャネル・ループ68及び68Aに、更に二次記憶ユニット70Aを接続したということである。二次記憶ユニット70Aは複数のテープ・ドライブ76を備えている。それらテープ・ドライブ76は、ディスク・ドライブ74に対する二次的な記憶装置であるといえる。

【0030】二次記憶ユニット70Aにディスク・ドライブ74とテープ・ドライブ76とを共に備えるようにすれば、光ファイバのループ上に構成された光ファイバ・チャネルの転送速度でディスク・ドライブからテープ・ドライブへ、またその逆へデータを直接転送することができるため、ホスト・サーバ62が余分な指令を発する必要がないという利点や、そのデータ転送作業と、ホスト・サーバ62が実行しているその他のアプリケーションとが、CPU時間、メモリ空間、或いはバス帯域幅に関して競合することがないという利点が得られる。

【0031】この図6に示した好適な実施例においては、一次記憶装置に装備したドライブはディスク・ドライブであり、二次記憶装置に装備したドライブはテープ・ドライブであるが、当業者には容易に理解されるように、記憶装置に装備するドライブはこれらに限られず、例えば、書き込み可能なレーザ・ディスクを使用した光ドライブや、RAMメモリ等のデバイスでもよく、更には、読み取り専用デバイスでよければCD-ROMを使用することもできる。以上を要約すると、この実施の形態は、一次記憶装置として用いる第1記憶手段と、その一次記憶装置に対する二次記憶装置即ちバックアップ用記憶装置として用いる第2記憶手段とを備えることを要件とするものである。

【0032】図7には、本発明にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの更に別の実施例が示されている。この実施例では、複数台のサーバにおいて互いに異なった種類のネットワーク・オペレーティング・システムが使用されている。この実施例は、マルチブル・ネットワーク・オペレーティング・システム共用記憶機構の構成を提供している。例えば、1つのサー

バ80はネットワーク・オペレーティング・システムとして「NT」を使用し、もう1つのサーバ82は「NetWare」を使用し、第3のサーバ82は「UNIX」を使用するということが可能である。この実施の形態にかかるアーキテクチャは、規格化された標準的な相互通接機構である光ファイバ・チャネルを用いて、互いに異なった種類のネットワーク・オペレーティング・システムを共用記憶ユニット70に接続するようにしたものである。

【0033】図8に示したのは、本発明にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの更に別の実施例である。この実施例では、複数台のサーバにおいて互いに異なった種類のサーバ・プラットフォームが使用されており、それらサーバ・プラットフォームは、図示例では「コンパック・プロリアント(Compaq Proliant)」、「HP 9000」、それに「IBM RS 6000」である。図7及び図8の構成は、種々のアプリケーションを実行するために様々なプラットフォームないしプロダクト・ラインを使用しなければならないユーザにとって特に有用なものである。例えば、証券金融業務用アプリケーションのうちにはサン・ワークステーション(Sun社のワークステーション)以外では動作しないものがある。そのようなアプリケーションを使用しているユーザは、ネットワーク・オペレーティング・システムとして「NT」を使用しているサーバを装備したシステムや、コンパック(Compaq)社製のサーバを装備したシステムへは、データを移送することができなかった。しかるに、本発明にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャによれば、光ファイバ・チャネル・ループを使用しているため、共用記憶装置に格納してある同じデータベースを、サン・ワークステーションとそれ以外のプラットフォームとで共用することができる。以上を要約すると、ここでは複数の種類のサーバ・プラットフォームが、この光ファイバ・チャネル接続式記憶機構アーキテクチャの接続機構を介して同じ記憶ユニットを共用することが可能とされている。

【0034】尚、図7及び図8に示した実施例の光ファイバ・チャネル接続式記憶機構アーキテクチャに、先に説明した幾つかの実施例にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャの特徴を組み込むことも可能である。例えば、完全冗長性を備えたアーキテクチャとすること、光ファイバ・チャネル・ループをカスケード接続したこと、複数台のサーバを接続したこと、複数台の記憶装置を接続したこと、二次記憶装置を備えたものとすること、ホット・プラギング機能を備えたものとすること、等々が可能である。

【0035】

【発明の効果】以上説明したように、本発明に係る光フ

アイバ・チャネル接続式記憶機構アーキテクチャは、光ファイバ・チャネル・ループを介して、複数台のサーバ（それらサーバのプラットフォームは同種のものでなくともよい）が、複数台の記憶装置に格納されているデータを共用できるようにしたものである。また、そのアーキテクチャは、完全冗長性、拡張性、ホット・プラギング機能等を備えたものとすることができる、二次記憶機構を装備したものとすることも可能である。本発明にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャによれば、光ファイバ・チャネル・ループに接続する種々の装置ないしデバイスを広く分散させて設置することができ、それが可能であるのは、光ファイバ・チャネルで構成するバスは、10キロメートルもの長さにすることができるからである。更に、光ファイバ・チャネルから成る相互接続機構のデータ・レートは、1ギガバイト／秒以上の高速にことができる。また、使用するケーブルも、従来のパラレル・バスを使用する場合と比べて非常に取り回しの容易なものとなり、それが可能であるのは、本発明にかかるアーキテクチャに使用しているバスが光ファイバを用いたシリアル・バスだからである。基本的に、本発明にかかるアーキテクチャを採用することによって、最近の大部分のマルチブル・サーバ・システムの必要記憶容量をまかなえるだけの十分に大きな記憶容量を有する、高速の大容量記憶システムが得られる。

【0036】以上に本発明の好適な実施の形態を幾つかを図示して説明したが、当業者には容易に理解されるよう、それら実施の形態に対しては、本発明の原理及び概念から逸脱すことなく種々の変更を加えることができる、本発明の範囲は請求の範囲に明記したとおりである。

#### 【図面の簡単な説明】

【図1】1台のサーバを1台の記憶装置に接続している従来のSCSIバスのブロック図である。

【図2】本発明の第1の実施例にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャのブロック図である。

【図3】本発明の第2の実施例にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャのブロック図である。

【図4】本発明の第3の実施の形態にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャのブロック図である。

【図5】本発明の第4の実施の形態にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャのブロック図である。

【図6】本発明の第5の実施の形態にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャのブロック図である。

【図7】本発明の第6の実施の形態にかかる光ファイバ

- ・チャネル接続式記憶機構アーキテクチャのブロック図である。

【図8】本発明の第7の実施の形態にかかる光ファイバ・チャネル接続式記憶機構アーキテクチャのブロック図である。

### 【符号の説明】

## 5.0 光ファイバ・チャネル接続式記憶機構アーキテクチャ

52 サーバ

## 5.4 光ファイバ・モキネル・ループ

\*10

\* 56 記憶装置

62 サーバ

## 64、64A ホスト・バス・アダプタ

66 メディア・モジュール

## 68、68A 光ファイバ・チャネル・ループ

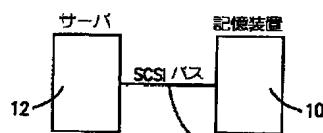
## 70、70A 記憶ユニット（記憶装置）

## 72、72A アレイ・コントローラ

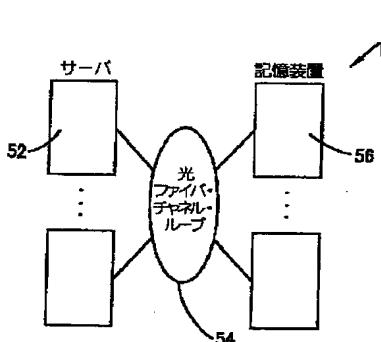
74 ディスク・ドライブ

76 テープ・ドライブ

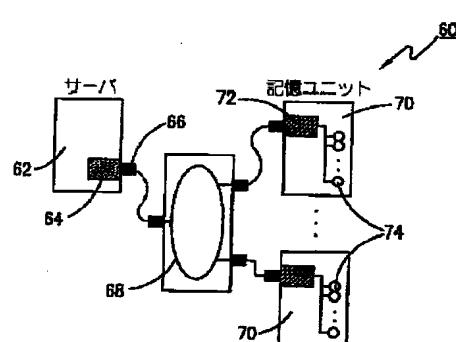
〔圖1〕



[圖2]

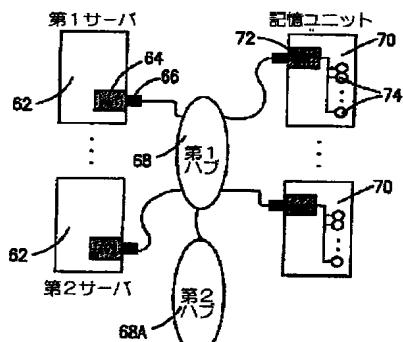


[圖3]

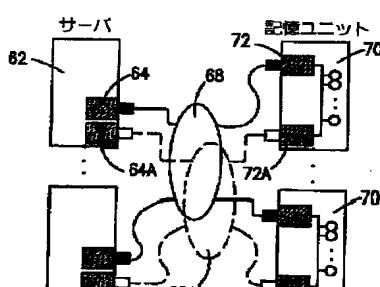


- ホスト・バス・アダプタ
- メディア・モジュール
- アレイ・コントローラ
- テイスク・ドライブ

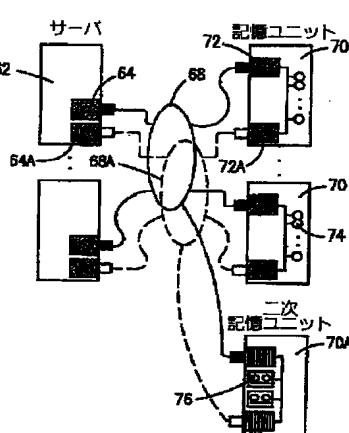
[図4]



[図5]



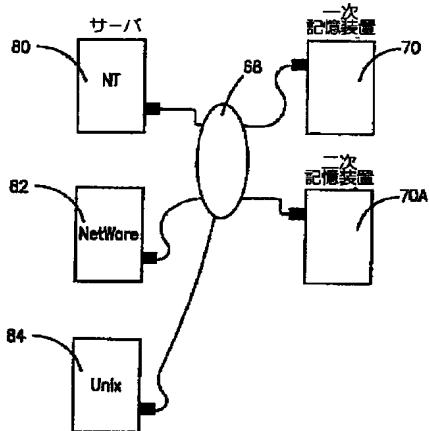
[圖 6-1]



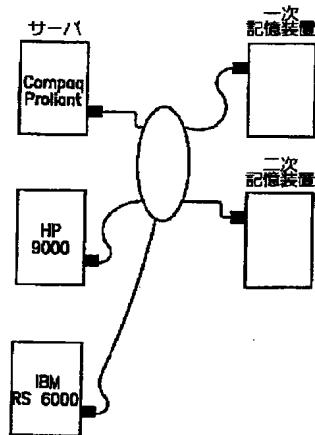
■ ホスト・バス・アダプタ  
■ アレイ・コントローラ  
— 第1経路  
— 第2経路  
□ ディスク・ドライブ

- ホスト・バス・アダプタ
- アレイ・コントローラ
- テープ・コントローラ
- ディスク・トラフィック
- テープ・トラフィック
- ディスク・ドライブ
- テープ・ドライブ

【図7】



【図8】




---

フロントページの続き

(71)出願人 591030868  
20555 State Highway  
249, Houston, Texas  
77070, United States of America  
(72)発明者 スティーブン・エム・シュルツ  
アメリカ合衆国テキサス州77069, ヒューストン, チェリー・ヒルズ・ロード 6923  
(72)発明者 ジョージ・ジェイ・スクルハマー  
アメリカ合衆国テキサス州77381, ザ・ウッドランズ, ビレッジ・ノール・プレイス 43  
(72)発明者 マーク・ジェイ・トンプソン  
アメリカ合衆国テキサス州77379, スプリング, バイン・ディストル・レーン 8206  
(72)発明者 トッド・ディー・ラッシュトン  
アメリカ合衆国テキサス州77375, トムボール, ウィロー・ブランチ・レーン 22518  
(72)発明者 ユージーン・イー・フリーマン  
アメリカ合衆国テキサス州77381, ザ・ウッドランズ, オッター・ボンド・プレイス

(72)発明者 デニス・ジェイ・アレクサンダー  
アメリカ合衆国テキサス州77379, スプリング, マホガニー・フォーレスト・ドライブ 18006  
(72)発明者 ライアン・エイ・カリソン  
アメリカ合衆国テキサス州77379, スプリング, パインウッド・フォーレスト 17519  
(72)発明者 トーマス・ダブリュー・グリーフ  
アメリカ合衆国テキサス州77379, スプリング, リップリング・ホール 6502  
(72)発明者 マイケル・エル・サボッタ  
アメリカ合衆国テキサス州77429, サイブレス, シーダー・ポイント・ドライブ 13506  
(72)発明者 デーヴィッド・エル・グラント  
アメリカ合衆国テキサス州77068, ヒューストン, フォーリング・クリーク 14919  
(72)発明者 ジェームズ・エフ・マカーティ  
アメリカ合衆国テキサス州77379, スプリング, ワンズワース・ドライブ 9327  
(72)発明者 ランディー・ディー・シュネイダー  
アメリカ合衆国テキサス州77379, スプリング, タワーストーン・ドライブ 9214  
(72)発明者 ウィリアム・シー・ギャロウェイ  
アメリカ合衆国テキサス州77083, ヒューストン, メシタ・ドライブ 14902